

USER CONTROLLED AUDIO QUALITY FOR  
VOICE-OVER-IP TELEPHONY SYSTEMS

BACKGROUND

Public Switched Telephone Network (PSTN) telephony offers only one level of service based on the use of a single audio quality regime. For example, the PSTN uses 64 kilobits Per Second (BPS) Pulse Code Modulated (PCM) audio with a 300-3400 Hertz passband. PSTN telephony further offers only one control regime that deterministically either grants or denies service rather than degrading service when a failure or network congestion occurs.

Voice-Over-Internet Protocol (VoIP) systems are often built to mimic this service regime but are also inherently more flexible. Certain VoIP protocols and algorithms can adapt to varying service levels. Example adaptations include either denying service or degrading service when resources cannot be reserved, or reserved resources become unavailable. Different encoder/decoders (Codecs) are switched in and out if bandwidth becomes more or less scarce. Forward Error Correction (FEC) can be added or removed as packet error rates change. Packetization intervals can also be varied to either limit bandwidth utilization or limit packet transmission rate.

To date these adaptations have been statically provisioned by the network designer and are not user controllable. The problem is that these adaptations may not be optimized for the current telephone users or the current VoIP call. For example, a given quality of service may be more than adequate for one user. However, that same given quality of service may be unsatisfactory to a different user.

5 The present invention addresses this and other problems associated with the prior art.

### SUMMARY OF THE INVENTION

A call adaptation system tracks adaptation schemes used for transmitting audio packets in a Voice Over IP call. A user response to the Voice Over IP (VoIP) call is monitored. The call adaptation system then dynamically varies the adaptation schemes used  
10 for transmitting the audio packets according to the monitored user response.

### BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagram of a communications network including a call adaptation system.

FIG. 2 is a detailed diagram of the call adaptation system shown in FIG. 1.

15 FIG. 3 is a block diagram showing how the call adaptation system modifies different adaptation parameters.

FIG. 4 is a block diagram showing how the call adaptation system modifies different adaptation parameters according to network congestion measurements.

20 FIG. 5 a block diagram showing how the call adaptation system operates in a Plain Old Telephone System.

FIG. 6 a block diagram showing the call adaptation system is implemented with a graphical user interface.

### DETAILED DESCRIPTION

25 FIG. 1 shows a communications network 12 that includes a packet switched network 26. The packet switched network 26 includes multiple network processing devices 28 that connect Voice Over Internet Protocol (VoIP) calls between different telephony end points. For example, a VoIP call 22 is established between endpoint 14 and endpoint 30. The

5 network processing devices 28 can be any routers, switches, gateways, concentrators, etc.,  
used for transferring packets over the packet switched network 26.

A call adaptation system 16 includes a user input device 18 that receives inputs  
directly from a user 32. The call adaptation system 16 allows the user 32 to dynamically  
control audio quality by dynamically varying call adaptation parameters used for conducting  
10 the VoIP call 22.

The user 32 initiates the VoIP call over the telephone endpoint 14. The user 32 then  
communicates with endpoint 30 through handset 34. If there is a desire to vary the perceived  
audio quality of the VoIP call 22, the user 32 adjusts input device 18. For example, the user  
32 may perceive the VoIP call 22 as having unsatisfactory sound quality. The user can turn  
15 input device 18 in a counter-clockwise direction in an attempt to increase sound quality of the  
VoIP call 22. The call adaptation system 16 then varies call adaptation parameters that  
improve sound quality.

In another example, the user 32 may think the sound quality of the VoIP call 22 is just  
fine. However, the user 32 wants to reduce the cost of the VoIP call 22. The user 32 turns  
20 the input device 18 in a clockwise direction in an attempt to reduce call cost. The call  
adaptation system 16 then modifies the call adaptation parameters, such as used bandwidth,  
to reduce the cost of the VoIP call 22. These adjustments to the adaptation parameters may  
reduce the sound quality of the VoIP call 22.

The user 32 can control how hard the call adaptation system 16 tries to vary the  
25 current sound quality in the VoIP call 22. The user 32 can control the adaptation parameters  
both before and during the VoIP call 22.

The input device 18 differs depending on the instrument used for establishing the  
VoIP call 22. For example, a physical knob or screen control may be used on an IP phone. A

graphical slider or similar icon representation may be used on a Personal Computer (PC) phone. The input device 18 can have two extreme settings and any number of intermediate settings. Examples of some of these embodiments are shown below in FIGS. 5 and 6.

FIG. 2 is a detailed diagram of the telephony endpoint 14 and the call adaptation system 16 shown in FIG. 1. A phone 36 generates audio signals 44 that are encoded into digital data by a voice encoder 38. A packetizer 40 converts the digital data into Voice over IP (VoIP) packets that are then transmitted over the packet switched network 26 through a transmitter interface 42.

The voice encoder 38, packetizer 40 and transmitter 42 use adaptation parameters 44 for encoding and packetizing the audio signals from phone 36. For example, different voice coder algorithms 46 are used by voice encoder 38 to convert the analog audio signals 44 into digital data. Two examples of voice coders are International Telecommunications Union (ITU) Standards G.723.1 and G.711. These different voice coders 46 encode different amounts of audio signals into digital data and ultimately vary the sound quality of the VoIP call 22.

In another example, different packet payload sizes 48 can be used by packetizer 40 when converting the digital data from voice coder 38 into VoIP packets. These packet payload sizes 48 can vary the amount of delay experienced during the VoIP call 22. Varying packet payload according to measured network congestion is described in copending U.S. Patent Application, Ser. No. 09/181,947, filed October 28, 1998, entitled: CODEC-INDEPENDENT TECHNIQUE FOR MODULATING BANDWIDTH IN PACKET NETWORK.

Another adaptation parameter is the type of Forward Error Correction (FEC) 50 used, if any, to prevent packet loss during the VOIP call 22. A bandwidth reservation parameter 52

5 identifies an amount of bandwidth, if any, the endpoint 20 attempts to reserve on the packet switched network 26 for the VoIP call 22.

A call adaptation controller 54 in the call adaptation system 16 adjusts the adaptation parameters 44 according to the user response detected from input device 18. The call adaptation controller 54 can also vary the adaptation parameters 44 according to the  
10 congestion measurements 24 received from packet switched network 26. This is described below in FIG. 4.

It should be understood that this is just an example of the types of adaptation parameters 44 that may be used to adjust the sound quality of the VoIP call 22. The call adaptation system 16 can use other parameters that vary the sound quality of the VoIP call  
15 22.

#### Mid-Call Resource Reservation

FIG. 3 shows in further detail how the call adaptation system 16 isolates the user 32 from network congestion by initiating or terminating resource reservation and/or by  
20 modulating packetization, forward error correction, and the codec. The call adaptation system 16 decides mid-call whether to reserve network resources for the VoIP call or to use a best effort scheme according to the monitored user response. For example, at a low sound quality user setting, the VoIP call may use a best effort scheme for establishing and conducting the VoIP call. At a higher sound quality user setting, the VoIP call may try to  
25 reserve network resources.

User response to the VoIP call is continuously monitored in block 56. If a low sound quality request is detected in block 58, the call adaptation controller 54 (FIG. 2) uses best effort packet traffic to establish or continue the current VoIP call. Best effort packet traffic is

5 defined as VoIP packets that are transmitted with no pre-reservation of network resources.

Thus, if a network processing device in the packet based network is congested, some or all of the VoIP packets may be dropped before reaching their destination point.

With the low sound quality request, a low quality voice coder is selected by the call adaptation system. The low quality voice coder uses a relatively small number of digital bits  
10 to represent a given portion of an audio signal. At the low sound quality setting no Forward Error Correction (FEC) and relatively long packet payloads might be used in the VoIP call.

If the user makes a first intermediate sound quality request in block 60, the adaptation system initiates a request to reserve network resources but also accepts best effort traffic while trying to reserve the network resources. One protocol used for reserving network  
15 resources is the Resource Reservation Protocol (RSVP). The RSVP protocol works through a protocol exchange which visits each network processing node along the VoIP media path. The RSVP messages request network processing devices along a media path to reserve a particular packet bandwidth.

20 The call adaptation system provides a unique resource reservation scheme that launches resource reservation mid-call. The first intermediate voice quality request detected in block 60 launches the RSVP reservation request mid-call in block 64. If the RSVP reservation request is accepted by all the network processing nodes in the media path, then the call adaptation system has locked down the resources for providing higher voice quality.  
25 If the user then requests even higher sound quality, the call adaptation system can increase the packet rate and be assured the network can handle the increased bandwidth requirement for the VoIP call.

5 If the RSVP request is not accepted, then the call adaptation system may not be able to improve voice quality if the user requests improved sound quality. In this case, there may be no improvement in sound quality if the user increases the voice quality request. The call adaptation system in block 64 may also switch to a better quality voice coder and/or reduce the packet payload size.

10 If a second intermediate sound quality request is detected in block 66, best effort traffic may still be used but a better quality voice coder and/or even shorter packet payloads may be used in block 68. In combination with the other adaptation parameter modifications, or by itself, FEC may be added to the VoIP call.

15 If a high sound quality request is detected in block 70, the VoIP call is conducted using only reserved network resources in block 72. A highest quality voice coder is used along with possibly an even shorter packet payload and FEC. The call adaptation system then continues to monitor for new user sound quality requests in block 56.

20 The example described above is for illustrative purposes. As described above, alternative implementations can also be used. For example, each user setting may include modification of a different one or different combinations of the adaptation parameters described above. Each setting may modify just one or a subset of the adaptation parameters described above. Alternatively, additional settings may provide smaller changes to the adaptation parameters or change adaptation parameters not mentioned above.

#### 25 Modifying Adaptation Parameters Based On Congestion Measurements

Referring briefly back to FIG. 1, to improve voice quality, it may be tempting to simply "crank up the gain" by injecting a higher packet rate into the VoIP call 22. However, this can actually increase congestion in the packet network 26 and be counter-productive to

5 improving audio quality. The call adaptation system 16 takes congestion measurements 24 into account through jitter and/or packet loss measurements. These congestion measurements 24 are conveyed back to the call adaptation system 16 using existing control protocols such as the Real Time Control Protocol (RTCP). The call adaptation system 16 takes the network congestion measurements 24 into account when adjusting the adaptation parameters to  
10 improve responsiveness to the user sound quality requests.

FIG. 4 shows how the call adaptation system 16 uses the congestion measurements 24 to determine what adaptation parameters 36 are used for varying the quality of the VoIP call 22. Referring to both FIG. 1 and FIG. 4, congestion measurements 22 are monitored in block 61 and the user response to the current VoIP call is monitored in block 63. In one example,  
15 the call adaptation system 16 may receive an RTCP message in block 59 indicating that a certain number of the VoIP packets are being dropped during transmission. A user response may also be detected in block 59 requesting an increase in sound quality.

Block 59 determines if the current congestion measurements or the current user response requires modification of the VoIP call adaptation parameters. For example, the  
20 number of lost packets identified in the congestion measurements may indicate that the sound quality requested from the current user response is not currently being met. Alternatively, a new user sound quality request may not be provided by the currently used adaptation parameters.

If it is determined that adaptation parameters need to be modified, the call adaptation  
25 system selects one or more adaptation parameters in block 65. The selected adaptation parameters are compared to the congestion measurements and user response in block 67. If the selected adaptation parameters do not provide the best chance of obtaining the sound



5     quality for the current user response with the current congestion conditions, the adaptation parameters are reselected in block 65.

For example, a voice coder may be selected in block 65 that increase the number of packets generated each second. However, there may currently be a large packet loss rate in the VoIP call. The call adaptation system may then opt to add Forward Error Correction  
10     (FEC) to the VoIP call instead of increasing the bit rate of the voice coder.

The selected adaptation parameters are modified according to the monitored user response in block 69. The call adaptation system then continues to monitor congestion in block 61 and the user response in block 63. If adaptation parameters still require  
15     modification, the system repeats the process described above. For example, FEC may not reduce the packet loss identified in the congestion measurements, the call adaptation system 16 may then opt to use a higher quality coder and/or vary packet payload size.

The call adaptation system uses linear regression based on current congestion measurements to find the optimal adaptation points for each of the adaptation parameters. The congestion measurements ensure that an adaptation envelope stays within acceptable  
20     congestion bounds. Such bounds are usually established through either the absolute packet loss rate or a first derivative of the absolute packet loss rate.

For example, if the packet loss rate is increasing, the call adaptation system adapts by cutting down on total consumed bandwidth. This linear regression is conducted in block 59 of FIG. 4 where the adaptation parameters are modified according to the monitored  
25     congestion measurements.

FIG. 5 shows another embodiment how Dual Tone Multiple Frequency (DTMF) signals are used to detect the user response to the VoIP call. A phone 72 is connected to a standard Plain Old Telephone Service (POTS) line 73. A user of phone 72 calls up a gateway

5 78 through a Central Office (CO) switch 76. The gateway 78 establishes a VoIP call 88 over a packet switched network 86 to another phone 90. The gateway 78 converts analog voice signals from phone 72 into VoIP packets.

The user of telephone 72 may wish to modify the sound quality of the VoIP call 88 as described above. The user generates a hook flash in phone 72 to first indicate a mode change  
10 to call adaptation controller 84. The user of phone 72 then presses buttons on the phone 72 to generate the DTMF signals 74. The pressed buttons identify the user response. For example, the "1" button by indicate a first sound quality request, the "2" button may indicate a higher sound quality request, etc. The CO switch 76 forwards the DTMF signals 74 to gateway 78. The call adaptation controller 84 detects the DTMF signals 84 and modifies the adaptation  
15 parameters 82 in an attempt to conform the quality of the VoIP call 88 to the user response represented by the DTMF signals 74. A Digital Signal Processor (DSP) 80 then uses the adaptation parameters 82 selected by call adaptation controller 84 to encode and transmit packets in VoIP call 88.

### 20 Graphical User Interface

FIG. 6 shows a computer 91 that includes a computer screen 92 and a keyboard 104. The computer 91 includes an audio communication device, such as two-way speaker 106 that allows a user to conduct a VoIP call. The user initiates the call by either dialing up a phone number or entering an Internet address of another VoIP endpoint. After the call is  
25 established, the user communicates using the two-way speaker 106, a headset, or alternative communication device.

During the VoIP call, the user can request the call adaptation controller 108 in computer 91 to vary sound quality of the VoIP call. A slider 96 is displayed on the screen 92

5 and allows the user to request changes in the sound quality. The user, either through a mouse 102 or the keyboard 104, moves a pointer 100 on slider 96. When pointer 100 is moved all the way to the left, the call adaptation controller 108 is notified that the user wants the highest sound quality. The user moves pointer 100 all the way to the right to request the lowest sound quality for the VoIP call.

10 In an Internet Service Provider environment, the service provider can offer different billing regimes for differing VoIP quality. The user can use inputs to the call adaptation controller 108 to see or control how much is being charged either before or during the VoIP call.

15 For example, the computer 91 also displays a slider 94 on screen 92 that identifies the cost of the VoIP call. The cost of the VoIP call may vary according to the used bandwidth, time of day, location of the VoIP endpoints, reservation of network resources, assigned call priority, etc. In one example, the cost is displayed on a cost per minute basis.

20 The user sees real time the result of requesting higher sound quality by viewing the cost slider 94. If the user is satisfied with the current sound quality, but would like to reduce call cost, the user can move pointer 100 to the right. If sound quality is inferior and the user does not object to increasing the call cost, the pointer 100 can be moved to the left.

25 The user can also move pointer 98 on slider 94 to limit the cost of the VoIP call. The pointer 98 is moved to the maximum amount that the user is willing to pay for the VoIP call. The call adaptation controller 108 then modifies the adaptation parameters to try and get the best possible sound quality for the selected call cost.

The devices used in the system described above can use dedicated processor systems, micro controllers, programmable logic devices, or microprocessors that perform some or all

of the adaptation and voice processing operations. Some of the operations described above may be implemented in software and other operations may be implemented in hardware.

For the sake of convenience, the operations are described as various interconnected functional blocks or distinct software modules. This is not necessary, however, and there may be cases where these functional blocks or modules are equivalently aggregated into a single logic device, program or operation with unclear boundaries. In any event, the functional blocks and software modules or features of the flexible interface can be implemented by themselves, or in combination with other operations in either hardware or software.

5        Having described and illustrated the principles of the invention in a preferred embodiment thereof, it should be apparent that the invention may be modified in arrangement and detail without departing from such principles. I claim all modifications and variation coming within the spirit and scope of the following claims.